DOCUMENT RESUME

ED 380 995 FL 022 818

AUTHOR Prideaux, Gary D.; Hogan, John T.
TITLE The Role of Markedness in Coding Episode Boundaries: Evidence from Oral and Written Discourse.
SPONS AGENCY Canadian Social Science and Humanities Research Council, Ottawa (Ontario).
PUB DATE [94]
CONTRACT 410-90-0109; 410-90-0125
NOTE 11p.; Truncated version of paper presented at a Conference on Organization in Discourse (Turku, Finland, 1994).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Contrastive Linguistics; *Discourse Analysis; English; Fiction; Foreign Countries; *Language Patterns; Language Research; *Linguistic Theory; Mandarin Chinese; *Oral Language; Programming (Broadcast); Research Methodology; Television; *Written Language
IDENTIFIERS *Markedness

ABSTRACT

An analysis of oral and written discourse focuses on the hypothesis that marked structures serve to code discourse boundaries and signal the advent of new discourse units. The hypothesis is examined against three independent sets of data: (1) discourse from a larger study, in which subjects provided oral or written narrative descriptions, in English, of the same event (from a movie); (2) an identical study except that subjects provided descriptions in Mandarin Chinese; and (3) analyses of two texts, both novels in English. In each analysis, markedness distribution in initial and non-initial positions was examined. It is concluded that all three studies provide independent supporting evidence for the hypothesis from both written and oral narrative discourse. Contains 18 references. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 380 995

THE ROLE OF MARKEDNESS IN CODING EPISODE BOUNDARIES: EVIDENCE FROM ORAL AND WRITTEN DISCOURSE

Gary D. Prideaux
John T. Hogan

Department of Linguistics
The University of Alberta
Edmonton, Alberta, Canada

1. Introduction.

Originating within Prague School phonology, the notion of markedness has received considerable attention from linguists for decades. Markedness theory claims that when two (or more) alternative forms are available, such as alternative construction types, the *unmarked* member of the set will tend to have several consistent characteristics, including being the more prototypical, the more frequent, the one with the widest distribution, and the like. Beyond the strictly formal domain, unmarked forms have been argued to be less costly in terms of processing recourses (Givón, 1990) and acquired earlier (Slobin, 1973) than their marked counterparts.

More recently, attention has been directed toward the discourse functions of marked structures. Fox (1987) and Givón (1993), for example have suggested that one important function of certain marked structures is to signal the beginning of a new discourse unit. In her examination of the distribution of NPs and anaphoric pronouns within the context of text data from both English and Tagalog, Fox (1987) found that the marked members tended to be found at the beginning of a paragraph. In a similar text study dealing with *before/after* subordinate clauses, Prideaux (1989) found that it was only within the most "oral" of written texts that marked structures were significantly correlated with paragraph-initial position, suggesting that this particular function of markedness might tend to be more of an oral than a written strategy. One problem with using written texts and especially of examining paragraph boundaries, of course, is that the conventions for paragraphing vary greatly according to style.

Nevertheless, the two major factors isolated above, that marked structures should consume more processing resources than their unmarked counterparts, and that in some cases marked structures appear to introduce new discourse units, together provide the foundation for a potentially important empirical hypothesis. If marked structures do consume relatively more processing resources, both for the speaker in constructing a message, and for the hearer in building a mental representation of the evolving discourse, then the appearance of such structures, or better, the selection of such structures, would be expected to inhibit processing at a certain points, point where new directions in the evolving discourse or narrative could be initiated. The purpose of the present paper, therefore, is to examine the hypothesis that marked structures do serve to code discourse boundaries and signal the advent of new discourse units.

FL022818

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Gary Prideaux
John Hogan

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this docu-ment do not necessarily represent official OERI position or policy

The hypothesis will be examined against three independent sets of data. The first set is taken from a larger experimental study in which a variety of processes were explored. In this study, subjects provided either oral or written narrative descriptions of the same set of event (from a short movie clip), and these productions were then examined for a variety of factors. The second study was identical to the first, except that the subjects provided oral descriptions in Mandarin, thereby providing cross-linguistic data. The third study involved two text analyses. Here two written narratives were examined, with special attention to the advent of new discourse or narrative units. Each study will be discussed briefly, and then general conclusions will be offered as to viability of the hypothesis.

First, however, attention must be addressed to a fundamental question involved in all such studies, namely how do we determine the advent of a new discourse unit? Clearly, we cannot simply use the presence of a marked structure as our operational definition, since such a move would render the entire enterprise viciously circular. One obvious move, for written data, is to use the written convention of paragraph as an indication of a new discourse unit. This move, adopted by Fox (1987) and Prideaux (1989), is highly suspect, however, since the canons of the written language require, or permit, new paragraphs for a variety of reasons which may have nothing at all to do with the overall narrative or discourse structure of the text. As a first approximation, the operational definition of a paragraph boundary as a new discourse unit might be tentatively acceptable, but in general terms, it is very weak, in spite of some evidence that readers of a language do have a sense of where paragraph breaks might fall (see Brown & Yule, 1983, pp. 97ff).

The problem is compounded even more when oral data are examined. How are we to know when a shift in discourse can take place? While some speakers may feel relatively comfortable with an indication of where a discourse changes direction, what we need is some operational definition of this shift in order to make empirical sense of the hypothesis under question. This question will be addressed for both the oral and written data under consideration below. It is important, however, that the question be exposed at the outset as one which must be clearly addressed if any sense is to be made of the hypothesis.

In section 2 below, the experimental data from English are reviewed, while in section 3, the Mandarin data are examined. In section 4, the text counts are presented, and finally in section 5, the overall conclusions are discussed.

2. The English Data.

The experiments reported here are part of a larger study in which several hypotheses concerning language processing were explored (Prideaux, Hogan, & Stanford, 1992). The methodology was similar to that found in Chafe's (1980) "pear stories" studies. In particular, we asked participants to watch a short film clip, a segment from the movie *Adam's Rib*, after which the participants either (a) wrote a short narrative description of what they had seen or (b) provided an oral description, which was taped, to a friend and again to a stranger. Thirty-two participants (16 male and 16 female) provided written narratives, while 40 participants (20 male and 20 female) provided two oral narratives each (one to a friend and one to a stranger), yielding 80 oral narratives.

Each of the narratives, written and oral, was transcribed and coded for a variety of factors including clause types, subordinate clause position, clause length, types of given

and new information, etc. For our present purposes, attention is restricted to those subordinate clauses or phrases initiated by the subordinate conjunctions as *before*, *after*, *when*, *while*, *since*, and *as*. Such complex sentences can either have the main clause first, followed by the subordinate clause (MC+SC) or they can have the subordinate clause first (SC+MC). It has been argued (by, e.g., Clark & Clark, 1977; Givón, 1983; *inter alia*) that the order MC+SC is the unmarked order, while the SC+MC order is marked:

1a. It started to rain before we could get to the car. (MC+SC, unmarked)

b. Before we could get to the car, it started to rain. (SC+MC, marked).

All such subordinate clauses were tabulated in all the oral and written narratives. However, in order to test the hypothesis that marked structures tend to initiate discourse units, it was necessary to find an independent measure of such junctures. This was done by having independent judges examine the film segment which served as the common stimulus for all the participants in order to construct a fine-grained analysis of all the actual episodes and events contained in the scene. The resulting analysis yielded some nine separate episodes, with each episode containing one or more events (see Prideaux & Hogan, 1993 for a detailed discussion). This analysis therefore provided an independent measure against which the narratives could be assessed. It was found, not surprisingly, that many of the participants did not mention some or even many of the events or episodes, while all mentioned a couple of the more salient. To determine whether a particular structure fell at the beginning or within an episode, we simply asked whether or not each subordinate clause was used as the first mention of an event in a given episode or not. The data from all the oral narratives, having first been examined to see if there were any major differences among them in terms of such factors as verbosity, types of structures used, etc., were pooled. The results for the oral narratives, tabulated as to position and markedness, are found in Table 1, while the results for the written narratives are found in Table 2.

Table 1. Markedness Distribution (Oral Narratives)

	Episode Position	
Structure	Initial	Non-initial
MC+SC (unmarked)	2	35
SC+MC (marked)	23	11

Table 2. Markedness Distribution (Written Narratives)

	Episode Position	
Structure	Initial	Non-initial
MC+SC (unmarked)	5	43
SC+MC (marked)	26	10

These results provide strong support for the hypothesis. For the oral data, a chi-square test reveals that marked structures are significantly more frequent at the beginning of an episode, while the unmarked structures are significantly more frequent within an episode ($\chi^2 = 30.1$, $p < .0001$). Similarly, for the written data, a chi-square test reveals exactly the same pattern ($\chi^2 = 33.7$, $p < .0001$). While there is some noise in the results, and in particular in the lower right hand quadrant (where several instances of marked

structures were found to be within episodes rather than at episode-initial positions), there is nevertheless strong support for the hypothesis. The exceptions tend to be cases in which a marked structure is used to refer back to the beginning of an episode or used to emphasize a particularly salient event in the scene. We can conclude, then, that these oral and written data, collected under controlled conditions from a large number of English speakers, do provide support for the hypothesis that marked structures, at least those involving subordinate clauses, do serve to demarcate discourse boundaries.

3. The Mandarin Data

There is always, of course, the possibility that the results reported in the section above are a function exclusively of English: maybe English is just idiosyncratic in using marked subordinate adverbial clauses in this manner. In order to address the issue of whether or not the phenomenon is operative in another language, Pu and Prideaux (in press) carried out exactly the same study as reported above, but using native speakers of Mandarin. In Mandarin, subordinate clauses must precede main clauses in all cases. However, four types of Mandarin subordinate clauses can be identified, of which two are marked and two unmarked. These are listed below:

2a. SC+MC (SC subjectless, marked)
Anmo wan le yihou, tamen jiu jiaohuan weizhi
massage finish PERF after they just switch positions
"After (she) finished massaging (him), they switch."

b. SC+MC (MC subjectless, unmarked)
nude zai anmo shi, kaiwanxiao di da le nande yixia
woman PROG massage when joking ADV slap man once
"While the woman massages (the man), (she) slaps the man playfully."

c. SC+MC (both clauses subjectless; unmarked)
yibian ku, yibian shuo
while cry while talk
"While (she is) cry(ing), (she is) talk(ing)."

d. SC+MC (marked)
ranhou na nude hai zai ku shi, nande jiu zou chuqu le
the woman still PROG cry when man just walk out PERF
"When the woman is still crying, the man walks out."

As argued in Pu and Prideaux (in press), in Mandarin complex sentences the subject typically appears in the first clause, regardless of its status, and typically is omitted in the second. If, however, the subject is given (previously mentioned), it can be omitted in both clauses. Thus, (2b) and (2c) are unmarked forms, while (2a), violating this general pattern, is marked. Such structures as (2d) are directly preceded (rather than followed, as in (2a, b)) by adverbial expressions of time or contrast phrases, and as such are typically marked and initiate new units (see Li & Thompson, 1979).

In the Mandarin study, twelve volunteers provided oral narratives of the same scene viewed by the English subjects. All twelve narratives were taped and transcribed, and all the complex sentences containing a subordinate clause were tabulated as to type and position within an episode. Precisely the same criteria were used to determine whether or not a particular sentence was episode initial or medial as was used for the English study. The results of the Mandarin oral study are found in Table 3.

Table 3. Markedness Distribution (Mandarin Oral Narratives)

	Episode Position	
Structure	Initial	Non-initial
SC+MC (unmarked)	2	25
SC+MC (marked)	22	6

Again, a chi-square test was carried out on these data, and the results clearly support the hypothesis, with a very strong tendency for marked structures to initiate episodes, but unmarked structures to fall within episodes ($\chi^2 = 28.3$, $p < .0001$).

These results suggest that, under the same stimulus conditions, for two very different languages, there is a strong tendency for marked structures to initiate an episode and unmarked structures to be found within the episode.

4. Text Count Data

To this point, the data discussed have all been taken from experiments in which participants provided admittedly very casual and informal descriptions of a particular scene they had witnessed. The question at this point is whether marked structures are used to demarcate discourse boundaries only in oral (and very oral-like written) descriptions, or whether this is a phenomenon with a broader scope. One way to examine this question is to look at naturally occurring discourses and conversations. Work on this is in progress: and we are examining a series of conversations among a group of family members and friends which we transcribed from a TV series of a couple of decades ago entitled *An American Family*. However, while the transcriptions are now all completed, the analyses have not yet been carried out.

Another way to examine the hypothesis is to look at the other extreme of the formality spectrum and examine written texts. As discussed above, Fox (1987) and Prideaux (1989) examined a series of written texts and looked at the beginnings and medial parts of paragraphs in order to assess the hypothesis. However, also as discussed above, the paragraph itself is a notoriously slippery construct on which to base firm empirical evidence. For this reason, a different perspective was adopted for the examination of written narrative texts. While the beginning of a paragraph may or may not initiate a new discourse unit, the advent of a new chapter almost certainly will. Accordingly, two novels were examined, namely *The Falling Woman* (Murphy, 1986), a novel of 31 chapters and 287 pages, and *Shaman* (Gordon, 1993), a novel of 74 chapters and 564 pages. For each novel, the first sentence of each chapter (P1,S1), was assessed according to whether or not it was a marked structure of any kind. To serve as a control for these numbers, samples were taken from two other places in each chapter. For the first novel, the controls were the first sentence of the third paragraph (P3,S1) and the second sentence of the

fourth paragraph (P4,S2), while for the second novel the controls were the first sentence of the fourth paragraph (P4,S1) and the second sentence of the sixth paragraph (P6,S2). There is obviously nothing special about these last two positions: they were selected as random positions to be sampled systematically. In each of the three positions, then, the sentence was assessed as to markedness. Table 4 provides the results for *The Falling Woman*.

Table 4. Markedness Distribution for *The Falling Woman*

	Position		
Structure	Chapter Initial (Para 1, Sentence 1)	Chapter Medial (Para 3, Sentence 1)	Chapter Medial (Para 4, Sentence 2)
marked	17	8	6
unmarked	14	23	25

Chi-square tests for these data reveal that there is a far stronger tendency for a marked structure to appear in chapter initial position than either medially as the beginning of the third paragraph ($\chi^2 = 5.43$, $p < .05$)or chapter medially within the fourth paragraph ($\chi^2 = 8.36$, $p < .005$). However, the second and third positions did not differ significantly in terms of markedness; that is there was no stronger a tendency for the first sentence of the third paragraph to be a marked structure than there was for the second sentence of the fourth paragraph ($\chi^2 = 0.37$, $p = .55$). This last result adds further support to the observation that simply paragraph initial position is not necessarily a good indicator of a thematic or discourse shift.

Table 5 contains the results for the text counts taken from the second novel.

Table 5. Markedness Distribution for *Shaman*

	Position		
Structure	Chapter Initial (Para 1, Sentence 1)	Chapter Medial (Para 4, Sentence 1)	Chapter Medial (Para 6, Sentence 2)
marked	38	21	12
unmarked	36	53	62

Here too, it is obvious that the chapter initial position enjoys a special status, with more marked structures in this location than in either of the other two control positions. Chi-square tests were again carried out to compare the three positions. The chapter initial position (P1,S1) contains significantly more marked structures proportionally than does either the first control position (P4,S1), which is itself paragraph initial ($\chi^2 = 8.15$, $p < .005$) or the second control position (P6,S2) ($\chi^2 = 20.42$, $p < .0001$). However, when the two control positions are compared against each other, the result is non-significant ($\chi^2 = 3.16$, $p = .0717$). These results parallel those for the first novel and suggest strong support for the hypothesis that, under the operational definition of a topic change at the beginning of a chapter, more marked structures are used to cue that change than are used

within a particular discourse or narrative unit. The scattering of marked structures at the beginning of paragraphs within a chapter could, of course, also signal a topic shift; at this point no strong claims can be made either for or against that possibility in the absence of an independent measure of discourse unit. Moreover, it is both possible and even expected that the marked structures found within a chapter could have other functions than solely that of signaling a thematic shift. Functions such as highlighting, foregrounding some particularly salient bit of information, or even moving the reader back to the beginning of a unit from within it are all possible.

It is interesting to note that while there is a wide variety of possible marked structures in English, those found in both novels in chapter initial positions were overwhelmingly preposed adverbial phrases or clauses. Only a few other types of marked structures were found in any of the positions sampled in this study, namely a few passives and a scattering of initial questions. Clearly, the fronting of an adverbial phrase or clause can serve as a scene setting device as well as a link to prior shared information or activating certain pragmatic information. Thus, fronted adverbials tend to be potentially multifunctional in this sense, and as such are congruent with a variety of discourse functions.

These results are compatible with those reported by Virtanen (1992) in her study of the discourse functions of adverbial placement in English. For example, in an experiment in which she asked participants to indicate paragraph boundaries in written texts, she found that the paragraph boundaries which were selected corresponded to major textual shifts at a remarkable 94 to 100% of the time. The results of her analyses of narratives and, to a lesser extent of procedural texts, reveal a similar pattern.

5. Conclusions

The three studies under discussion have all, independently, provided supporting evidence for the hypothesis that one important function of marked structures is to code the onset of a new discourse (or narrative) unit. It is important in such studies to try to secure the phenomenon under consideration with data from a variety of sources, and for this reason, both oral and written descriptive narratives from English were bolstered by oral data from Mandarin and with more formal written data, again from English. In all these cases, for both languages, the phenomenon seems to shine through with some considerable robustness. It is almost as if the speakers and writers have adopted a kind of "text strategy" (in the sense of Virtanen, 1992; Enkvist, 1987) of using marked structures to code the major text junctures, in both the oral and written forms of the language(s).

A final observation will serve to close this contribution: Much work in discourse analysis tends to be based on a single source or on highly idiosyncratic sources (a single dialogue, a taped conversation on the telephone, etc.). While not wishing to suggest that such studies are not useful, we can nevertheless observe that controlled experiments involving many participants carrying out the same task do provide a relatively more secure data base. Similarly, extensive text counts from more than a single source, properly controlled and assessed statistically not only provide empirically secure data but also, as in the instances here, point to areas where the hypothesis fails, thereby generating questions as to what other factors might be at work. In the present case, we find support for the hypothesis, but we also find cases where marked structures are used at non-junctures,

thereby raising the obvious question as to what other functions such structures might have.

Finally, we would suggest that research within the domain of discourse analysis needs to pay more attention to those cognitive processing factors involved in the organization of discourse than is typically done. Again, while not to suggest that the familiar sociolinguistic factors (gender, power levels, age, social distance, etc.) typically cited in much discourse work are not important, we would observe that those processing effects deriving from our common cognitive architecture are also of considerable importance and are also deserving of our attention as we attempt to uncover those many factors which contribute to the organization and structure of discourse. This study has been an attempt in just that direction, an attempt to examine a discourse function which has its source in, and is constrained by, our cognitive faculties.

Note

1. The research reported here was supported in part by Social Sciences and Humanities Research Council of Canada Research Grants 410-90-0125 and 410-90-0109, for whose support we are grateful. A truncated version of this paper appears in the Proceedings of the *Organization in Discourse* Conference, Turku University, Turku, Finland.

References

Abraham, E. (1991). Why 'because'? The management of given/new information as a constraint on the selection of causal alternatives. *Text, 11*, 323-339.

Brown, G., & Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.

Chafe, W.L. (Ed.) (1980). *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood, NJ: Ablex.

Clark, H.H., & Clark, E.V. (1977). *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.

Enkvist, N.E. (1987). A note toward the definition of text strategy. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung, 40*,19-27.

Fox, B. (1987). Morpho-syntactic markedness and discourse structure. *Journal of Pragmatics, 11*, 359-375.

Givón, T. (Ed.) (1983). *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: Benjamins.

Givón, T. (1990). *Syntax: A Functional-typological introduction, Volume 2*. Amsterdam: John Benjamins.

Givón, T. (1993). Coherence in text, coherence in mind. *Pragmatics & Cognition, 1*, 171-227.

Gordon, N. (1993). *Shaman*. New York: Signet.

Li, C.N., & Thompson, S.A. (1979). Third person pronouns and zero anaphora in Chinese discourse. In T. Givón (Ed.), *Syntax and semantics, Volume 12: Discourse and syntax* (pp. 311-335). New York: Academic Press.

Murphy, P. (1986). *The falling woman*. New York: Tom Doherty Associates.

Prideaux, G.D. (1989). Text data as evidence for language processing principles: The grammar of ordered events. *Language Sciences, 11*, 27-42.

Prideaux, G.D., & Hogan, J.T. (1993). Markedness as a discourse management device: The role of alternative adverbial clause orders. *WORD, 44*, 397-411.

Prideaux, G.D., Hogan, J.T., & Stanford, L.M. (1992). *Syntactic correlates of cognitive and social factors in controlled discourse production*. Final Report, Social Sciences and Humanities Research Council of Canada Research Grant 410-90-0125.

Pu, M.M, & Prideaux, G.D. (in press). Coding episode boundaries with marked structures: Evidence from Mandarin. *Canadian Journal of Linguistics*.

Slobin, D.I. (1973). Cognitive prerequisites for the development of language. In C.A. Fergurson & D.I. Slobin (Eds.), *Studies of child language development* (pp. 175-208). New York: Holt, Rinehart & Winston.

Virtanen, T. (1992). *Discourse function of adverbial placement in English: Clause-initial adverbials of time and place in narratives and procedural place descriptions*. Turku/Åbo: Åbo Akademi University Press.

\papers\eric.doc
January 26, 1995